Available online at www.sciencedirect.com

ScienceDirect

Cognitive Systems Research 27 (2014) 1–10

Cognitive Systems Research

www.elsevier.com/locate/cogsys

The evolution of syntax: Signs, concatenation and embedding

Action editor: Rajiv Khosla

Erkki Luuk [a,*], Hendrik Luuk [b]

[a] *Institute of Computer Science, University of Tartu, Estonia*
[b] *Department of Physiology, University of Tartu, Estonia*

Received 5 December 2012; accepted 24 December 2012
Available online 17 January 2013

Abstract

The paper argues that the structure, derivation and evolution of syntax is given by the sequence (elements, concatenation, embedding). We discuss the implications of this sequence for language and the numeral system in general and on the evolution of language in particular. A four-stage model of the evolution of syntax, broadly compatible with several earlier scenarios, is proposed. The four stages are (1) signs, (2) increased number of signs, (3) commutative concatenation, and (4) noncommutative concatenation. We support the model by showing that its stages can be adaptive per se, which could explain why they evolved. We also identify two preconditions for maintaining the stages: stage (2) depends on the ability to conceptualize asymmetric relations between concepts and the adaptiveness of stage (3) depends on cultural constraints on linguistic interpretation.

Keywords: Language; Evolution; Syntax; Concatenation; Embedding; Sign; Numeral system; Compositionality; Semantics

1. Introduction

Although the **capacity** for language is part of our genetic endowment, language is, essentially, a technological innovation, and one that rather evolved to fit the brain than vice versa (Christiansen & Chater, 2008; Doumas & Hummel, 2005). In modelling language evolution, the following scenario is widely agreed upon:

preadaptations [1] → protolanguage (→ preadaptations [2]?) → syntactic language

Certain preadaptations [1] were necessary for protolanguage to emerge. Anatomical preadaptations included changes in the brain anatomy; anatomical preadaptations for speech included changes in middle and/or inner ear anatomy, an enhanced thoracic innervation and a re-configuration of the tongue and vocal tract (Boë et al., 2007; Fitch, 2000; MacLarnon & Hewitt, 1999; Martínez et al., 2004; Wynn, 1998). Depending on one's theoretical standpoint, cognitive preadaptations could have been, e.g., theory of mind and relational reinterpretation (Call & Tomasello, 2008; Penn, Holyoak, & Povinelli, 2008; Penn & Povinelli, 2007). As protolanguage is, essentially, a language without syntax, it refers to either a holophrastic or arbitrarily concatenated language. Although culturally downgraded, both of these variants are exceedingly common in natural communication, e.g. in ellipsis, simple dialogues and giving orders. In fact, sentences are frequently difficult to identify in spoken discourse (Bowie, 2008). Although there are substantial structural differences between protolanguage and syntactic language, the main

Abbreviations: CARC, conceptualization of asymmetric relations between concepts; CCLI, cultural constraints on linguistic interpretation; mya, million years ago; N, noun; P–H, 'pyow–hack'; V, verb

* Corresponding author. Address: Institute of Computer Science, University of Tartu, Postimaja pk 149, 51004 Tartu, Estonia. Tel.: +372 58896309.

E-mail address: erkkiluuk@gmail.com (E. Luuk).

1389-0417
http://dx.doi.org/10.1016/j.cogsys.2013.01.001

functional difference is that, in syntactic language, linguistic form constrains interpretation better than in protolanguage, otherwise the expressive powers of the two variants are comparable. For example, it has been proposed that the difference between protolanguage and syntactic language is roughly of the order of that between pidgin and creole (Bickerton, 1990; Givón, 1998). In any case, protolanguage would have been sufficient to support all these properly symbolic or symboling-dependent activities discussed in Section 2. As to why protolanguage was eventually substituted with syntactic language, the most plausible explanation is that the transition reduced ambiguity and facilitated interpretation. It is unknown whether it was a solely technological innovation or required some additional anatomical and cognitive preadaptations [2]. However, see Hauser, Chomsky, and Fitch (2002) and Chomsky (2010) for the proposal that the preadaptations included a neurally implemented recursion. In linguistics, there is a sharp difference between historical (up to 10 000 years) and evolutionary (10 000 to millions of years) timescales. There is no concept of 'languages' contiguous to present day natural languages for the evolutionary timescale. As protolanguage pertains to the evolutionary timescale, it is cross-linguistically universal by definition. In the following sections, we propose a novel, universal and parsimonious model of the evolution of syntax, substantiate it and show the adaptiveness of its stages.

2. The evolution of syntactic compositionality of language

Martin A. Nowak and colleagues have established a mathematical framework for modeling the evolution of language based on evolutionary game theory (Nowak, Komarova, & Niyogi, 2001; Nowak & Krakauer, 1999; Nowak, Plotkin, & Jansen, 2000). Nowak and Komarova speak of 'compound signals': "Word stems /—/ of human languages are elementary signals, but phrases, sentences or any syntactic structures in human languages represent compound signals" (Nowak & Komarova, 2001, p. 290).

Compound signals imply relations between the concepts that they refer to. In natural language, the generic principle for compound signals is asymmetric dependency[1] (head-dependent, stem-affix, modified-modifier, main-subordinate clause, etc.). Thus, the **conceptualization of asymmetric relations between concepts** (CARC) is a cognitive prerequisite for language. From the viewpoint of CARC, the following statements are equivalent: concept A depends on concept B, A is caused by B, A contains B, A includes B, B belongs to A, B is a part of A, etc. The simplest kind of representations we regard as concepts are secondary representations in the sense of Perner (1991): cross-modal mental models capable of representing past, future, or imaginary objects or events, or representing the representational content of other representational systems. According to Perner (1991, p. 7), secondary representations are distinct from (and intermediate between) primary representations and metarepresentations.

In addition to relating two concepts asymmetrically, CARC enables conceptual compositionality (e.g. father = male parent, $2 = 1 + 1$, etc.) and semantic embedding (explained in the next paragraph). The adaptivity of CARC lies in an increase in the ability to plan one's behavior owing to the conceptualization of asymmetric relations governing the physical world. The effects of CARC include the conceptualization of containment hierarchies of depth 2 and more, causality, definitions, the concepts of knowledge and ownership, etc. The possibly uniquely human **semantic synthesis** ability, proposed by Dessalles, is also an effect of CARC. In describing protolanguage, Dessalles (2008, p. 56) gives the following example of semantic synthesis: "Listeners must integrate the different associations triggered by the different words, 'stranger', 'plain', 'fire' into one single state of affairs, instead of imagining several disconnected situations". Not only syntactic (clauses) but also morphosyntactic (inflected words) and discourse pragmatic (discourse context) devices are compound signals that subsist on CARC. It should be noted though that while clause and discourse are almost always compound (imply semantic embedding), phrases and word forms are frequently elementary. Thus we have to discern at least these four levels of semantic embedding (cf. below).

However, a compound signal is not the first step towards syntax. **Concatenation** is necessarily a compound signal only from the viewpoint of modern syntax. A protosyntactic concatenation lacks at least two features characteristic of modern syntax: grammar and semantic embedding. We define **semantic embedding** as follows: a meaningful linguistic unit in another meaningful linguistic unit, e.g. a phrase in a phrase, a word in a phrase, a word in a sentence, a word in a discourse, a morpheme in a word, etc.[2] A protosyntactic concatenation of any two signals A and B (e.g., *dog run*) is not a compound signal (in this case, a putative sentence), just two signals concatenated. It is only after extensive (and eventually, grammatically constrained) use that AB becomes a compound signal. Diessel and Tomasello (2005) have found that initially children use verbs like *think, know, see* exclusively in first person, present tense, never negated. Instead of embedding, this seems to be concatenation of a fixed form (*I know/think/see*) with a sentence. Similarly, in Jackendoff's (1999, p. 273) model of steps in the evolution of language, concatenation precedes the "use of symbol position to convey basic semantic relationships", which implies grammar and embedding (cf. Dessalles, 2006; Johansson, 2006). Thus, we arrive at Table 1.

[1] Incidentally, this also provides clues as to language evolution and change. As a rule of thumb, the necessary components antedate their dependents (cf. Budd, 2006).

[2] By syntactic embedding is usually meant the subcase of semantic embedding that operates on phrases and sentences.

Table 1 shows the logical and temporal succession of stages of syntactic compositionality, [...] marks a precondition for maintaining the stage. The stages are ordered vertically with each stage describing a state of syntactic compositionality achieved (e.g., ‘concatenation of signs’). Table 1 is hierarchical, i.e. at each stage the conditions stipulated by the previous stages (above them) apply as well. This accords with the evolutionary principle of building on rather than expunging the earlier stages. The timing of the stages is relative, i.e. the intervals between them might not be equal. According to this scale, two major steps in the evolution of syntactic compositionality are 1) from isolated signs to concatenated signs and 2) from concatenated signs to embedded signs. ‘Signs’ refer to distinctly meaningful signs – probably symbols but this is not so clear for the earlier stages (1)–(3), which might have had predominantly iconic or indexical signs (e.g. in gestural or vocal-gestural modality – Bickerton, 2003; Steels, Kaplan, McIntyre, & Looveren, 2002). An increased number of signs is attested as a prerequisite of language and a payoff condition for compound signals (Christiansen & Kirby, 2003; Jackendoff, 1999; Nowak & Komarova, 2001).

We assume that the ability to conceptualize asymmetric relations between concepts is a precondition for maintaining stage (2) (increased number of signs). CARC is implied by the concepts that subsume asymmetric relations, e.g. ‘influence’, ‘cause’, ‘result’, ‘kill’, ‘throw’, ‘heal’, ‘eat’, etc. As signs for such concepts cannot appear before the ability to entertain the concepts themselves, CARC is a prerequisite for a vast number of signs. As predicates, these signs are more complex than simple arguments (*tree*, *man*, etc.) and one-place predicates (*sleep*, *run*, etc.), i.e. they could be evolutionarily later additions to the vocabulary (cf. Heine & Kuteva, 2007; Luuk, 2009). At the same time, CARC is not a **sufficient** condition for stage (2). Thus, though this is certainly plausible, it is not a priori clear that all species that are unable to attain stage (2) would lack CARC.

By free concatenation we mean commutative concatenation, i.e. concatenation of elements regardless of their succession. Cultural constraints on linguistic interpretation (CCLI) are required for stage (3), free concatenation, for the latter to contribute to the individuals’ fitness (see below and Sections 3.1, 4). For example, in *kill man rat*, did a rat kill a man or vice versa? Our present CCLI suggest that it was probably the man who did the killing. CCLI are automatically evoked in the contexts where interpretation is

Table 1
The evolution of syntactic compositionality of language.

(1) signs
↓
(2) increased number of signs [CARC]
↓
(3) free concatenation of signs [CCLI]
↓
(4) grammar ~ constraints on concatenation of signs ~ semantic embedding

linguistically highly underspecified (in modern language, the specification is done by grammar). A relative freedom of concatenation is implied by the second payoff condition for syntactic communication “the compound signals must be able to encode the relevant messages in such a way that individual components occur in many different messages” (Nowak & Komarova, 2001, p. 291). Nevertheless, the freedom must be constrained by interpretation, either by CCLI (as in protolanguage) or by CCLI and grammar, otherwise coherent communication cannot emerge (cf. Jackendoff, 1999; Nowak & Krakauer, 1999). Conversely, the need of grammar arises only if communication about many different events is required – a language must have more relevant sentences than words (Nowak & Krakauer, 1999; Nowak et al., 2000), which in turn presupposes a relative freedom of concatenation. CCLI are implicit, whereas grammar provides explicit constraints for linguistic interpretation. From this, one can conjecture that the need for grammar arises when CCLI become inadequate.[3] This condition is met under the following circumstances: large group sizes, high levels of intragroup diversity, a growing need for intergroup communication or intragroup specialization. It is easy to observe that all these parameters indicate social sophistication.

3. Background

3.1. Evolutionary context

One of the best proxies for culture is a recording of the group’s experience on an external storage. As the forms and meanings of (proto)linguistic signs are shared by convention, both language and protolanguage count as external storage. Of course, as compared to written language, spoken and signed languages are ephemeral external storage that depends more on memory. Hence the possible significance of rhythm and melody as additional mnemonic cues for (proto)language. Observe also that a combination of sound and gesture, as in normal face to face discourse, provides more mnemonic cues than the discourse that is either exclusively signed or spoken (as it exhibits signal redundancy – which partly explains our automatic tendency to gesture while talking). As protolanguage is an external storage, culture either antedates protolanguage or is contemporaneous with it. Theoretically, a prelinguistic external storage could have made use of non-linguistic symbols or non-symbolic signs (i.e. icons and indices). Observe that, unless they follow distinctive and elaborate styles, the shapes of functional artifacts (e.g. tools) are

[3] Hurford (2010, p.c.) referred to the theoretical alternative that language becomes decorated with grammar as a badge of group identity. This is less plausible because a badge of group identity can be virtually anything (from handaxe to handshake style) whereas explicit constraints for linguistic interpretation are a narrow superset of grammar. Moreover, as the explicit (both lexical and grammatical) constraints are usually unique, they provide a badge of group identity for free.

more parsimoniously interpreted as suboptimal solutions to the tool material vs. task problem than an external storage of group's experience (cf. Wynn, 1991, 2002). The first signs of culture in this sense are mode 2 tools from 1.65 mya[4] (Klein, 2000). Mode 2 tools appear within the time frame for the earliest circumstantial evidence for language (which, in all likelihood, was a protolanguage). This evidence includes *Homo erectus*' successful colonization of much of the Old World (from Africa and Western Europe to Java, China and, possibly, Central Siberia) and its adaptation for enhanced vocalizations as compared to australopithecines (Ascenzi, Benvenuti, & Segre, 1997; Asfaw et al., 2002; Bar-Yosef & Belfer-Cohenb, 2001; Gibbons, 1998; Larick et al., 2001; MacLarnon & Hewitt, 1999; Meyer, 2003; Meyer, Lordkipanidze, & Vekua, 2006; Waters, Forman, & Pierson, 1997). The evidence also indicates that, by 0.8 mya, *H. erectus* had crossed substantial stretches of open water (at least 19 km) in Indonesia (Morwood, O'Sullivan, Aziz, & Raza, 1998). In sum, the circumstantial evidence brackets the emergence of (proto)language between 0.8 and 2.3 mya. The latter date corresponds to the appearance of *Homo habilis*, the first known *Homo* species (Kimbel, Johanson, & Rak, 1997). As *H. habilis* is the direct ancestor of *H. erectus* (Spoor et al., 2007), and a species that was not scrutinized by MacLarnon and Hewitt (1999), it is possible that *H. habilis* was anatomically adapted to speech as well (see Tobias, 1998).

In natural language, grammar, constrained (i.e. noncommutative) concatenation of signs and semantic embedding are coextensive. Unless we are dealing with a purely phonological (e.g. Vowel First) constraint, noncommutative concatenation is an asymmetric relation between meaningful units (signs), which in turn entails semantic embedding. As any asymmetric relation between meaningful units A and B (usually a head-dependent relation) stipulates a higher-order meaningful unit A–B, we have semantic embedding (a meaningful unit in another meaningful unit). Conversely, semantic embedding entails two levels of meaningful units, the boundaries of which can be given (over serial channel) only by concatenation. Over serial channel, embedding entails concatenation (e.g. [B[A]B] subsumes concatenate [B + A + B]).

From what we know, a primitive grammar might have had any of the following rules: the noun/verb distinction, Agent First, Focus Last, grouping, and noun-noun compounds (Jackendoff, 1999). All these rules imply semantic embedding and noncommutative concatenation. In modern language, semantic embedding (or noncommutative concatenation, here marked by [$\cdots + \cdots$]) constitutes the levels of the following grammatical units[5]: word [*run* + *s*], phrase [*a* + *man*], and clause (both relative and main clause, e.g. [[*a* + *man*] + [*run* + *s*]]). It is possible to have multiple phrasal embedding, as in [[[[*John's*] + *mother's*] + *cat's*] + *tail*], and multiple clausal embedding, e.g. [*He met the writer* + [*that the man* + [*who was ill*] + *had seen before*]]. All these rules are stipulated by grammar. Of all the syntactic units (phrase, morphological word, relative and main clause), the sentence or main clause is communicatively the most fundamental. There is hardly any need for phrases before there is a sentence, and there is not much need for morphology before syntax. Although morphology is used to make semantic distinctions (one/many, male/female, etc.), its main function is to serve syntax in argument, predicate, and argument-predicate relation marking (Luuk, 2009). Thus, the first syntactic unit was probably functionally equivalent to a sentence. There would be more than one possibility for this. Given the availability of semantically diverse stem categories, the simplest solution would have been to concatenate arguments and predicates, as in [*man go*]. Alternatively, with a categorially uniform stem choice, a solution would have been to concatenate different semantic roles, as in [*man forest*], interpreted as 'man go to forest'. Due to the opacity of interpretation the second possibility seems less likely but, as the categorial contents of the set of input stems is not known, the more plausible scenario cannot be established with certainty.

The general principle of grammar is the head-dependent relation, i.e. the principle of asymmetric dependency. Thus, grammar and semantic embedding presuppose CARC. Grammar and semantic embedding are inconceivable without CARC, whereas the latter is perfectly conceivable without language, grammar and semantic embedding. As CARC is prelinguistically useful (e.g. in planning), there is a fair chance that it antedated language.

Interfaces to phonology and semantics aside (Hauser et al., 2002; Jackendoff, 2002; Nowak & Komarova, 2001), the three building blocks {signs, concatenation, embedding} are all that is required for syntax – any syntax can be built (and described) with them[6] – while some of them are redundant in describing pre-syntactic stages. Noncommutative concatenation of signs yields the head-dependent relation for free (see above). Observe that one cannot speak of natural language syntax until stage (4) is achieved. Natural language syntax is qualitatively different from the raw syntax of other species (e.g. birds) communication systems in being semantically compositional (Gardner, Naef, & Nottebohm, 2005; Hurford, 2004). Given the accounts that apes and dolphins can be trained to learn symbols and understand primitive sentences in captivity, a proficiency seemingly pertaining to at least stage (3), it is puzzling that, to the present knowledge at least, they have developed no stage (3) communication system in the wild (Herman, Richards, & Wolz,

[4] Mode 1, appearing 2.5 mya (Klein, 2000), is essentially an ape-grade technology (Wynn, 1991).

[5] Discourse is not a grammatical unit because there is a consensus (articulated by, e.g., Meillet and Bloomfield – Graffi, 2001, p. 1843) that the sentence is the largest unit of grammar.

[6] In some languages (e.g. Dutch) there is an intermediate condition between concatenation and embedding – cross-dependency (Christiansen & Chater, 2003). For instance, $N_1N_2V_1V_2$ is an example of cross-dependency between N_1V_1 and N_2V_2.

1984; Savage-Rumbaugh, Shanker, & Taylor, 1998). One explanation that has been proposed to this curious inaptitude to commune is a lack of motivation (Bickerton, 2003; Seyfarth, Cheney, & Bergman, 2005; Szamado & Szathmary, 2006). Indeed, the degree of communication that gets rewarded in human societies is much higher than that of among other primate species (Knight, 2002). Yet, there is no doubt that our ancestors were once in the same situation the other primates are now. Thus it is unlikely that a lack of motivation could be a sufficient explanation for all, as Bickerton puts it, "relatively large-brained species" (Bickerton, 2003, p. 83). On the more technical side, Nowak et al. have some other possible explanations (Nowak & Komarova, 2001). Certain conditions have to be met before natural selection can see the advantages of compounding: 1. The total number of relevant messages has to exceed a critical value, 2. The compound signals must be able to encode the relevant messages in such a way that individual components occur in many different messages. Plausibly, these conditions are not met by non-humans. But why? We hypothesize that the degree of differentiation of conceptual structure in non-humans is insufficient to support these conditions. Specifically, there seems to be something unique about the human capacity for hierarchical conceptualization (but it is difficult to tell what exactly – see Chomsky, 2010; Dessalles, 2008; Fauconnier & Turner, 2008; Hauser et al., 2002; Luuk & Luuk, 2008; Penn & Povinelli, 2007; Penn et al., 2008; Premack & Woodruff, 1978; Suddendorf & Corballis, 2007; Tomasello, Call, & Hare, 2003; Tulving, 2005, for different hypotheses, some of which appear to be already falsified – Correia, Dickinson, & Clayton, 2007; Osvath, 2009).

Curiously, only one possible example of a semantically compositional syntax, the extremely limited communication system of honeybees,[7] is found in non-humans in the wild, and no clear example of semantically compositional communication has been found in non-human vertebrates (Hurford, 2004; Michelsen, 1999; von Frisch & Lindauer, 1996). There are bird songs, cetacean songs, and primate 'long calls' built up out of smaller units, but the units are not meaningful on their own, and/or different combinations are not distinctively meaningful (Jackendoff, 1999; Ujhelyi, 1998). This argument applies also to the special case of putty-nose monkeys (Arnold & Zuberbuhler, 2006). These monkeys produce two calls, 'pyows' and 'hacks' in response to, mainly, leopards and eagles, respectively. The researches found that the monkeys also produce a third call, 'pyow–hack' (P–H), and observed that P–H triggers group movement. In addition, although the putty-nosed monkeys sometimes move through the canopy to escape from an approaching leopard, this strategy is avoided when threatened by large raptors, as it would increase the risk of attack. Leopard growls were played back to 17 different monkey groups. In nine groups, the male produced call series containing at least one P–H. The researchers found that, 20 min later, the groups whose males had produced P–H had traveled significantly farther than other groups. It is important to note that P–H is not a semantic combination, and complies with P and H due to loud call repertoire constraints only (Arnold, p.c.). Elementary signals must retain their reference within a compound for the latter to be a semantic combination. First, it is not evident that P and H have retained their reference in P–H. Second, the monkeys' behavior would seem irrational if P and H had retained their reference, as movement is avoided when threatened by large raptors (H), as it increases the risk of attack.

3.2. Comparison with earlier approaches

Although the evolution of syntax has been of considerable interest to researchers, there are surprisingly few explicit models. This section compares our model with those explicit models and/or general approaches that are more compatible with Table 1.

Bickerton (1998) subscribes to a scenario with stages (1), (3) and (4), omitting (2). His scenario is more general than Jackendoff (1999), which proposes a detailed model. The differences between Jackendoff's and our model are following. (a) Our model is more universal: where Jackendoff speaks of 'symbols', we have 'signs'; Jackendoff's stages "use of symbol positions to convey basic semantic relationships" and "hierarchical phrase structure" are subcases of semantic embedding, i.e. conflated in our stage (4). (b) In Jackendoff's model, there is no link between "use of an open, unlimited class of symbols" and "concatenation of symbols", corresponding to our stages (2) and (3)–(4) that are linked both evolutionarily and derivationally. (c) In his model, the distinction between commutative and noncommutative concatenation is implicit rather than explicit. Nowak et al. (Nowak, 2000; Nowak & Krakauer, 1999; Nowak et al., 2000, 2001) do not analyze language evolution into an explicit succession of stages. However, the following stages can be inferred: phoneme-object pairs (1), increased number of words (2), grammar (the word types N and V) (4). As such, their model omits stage (3). Notice also the difference between 'phoneme-object pair' and 'word' – not all words are phoneme-object pairs (both are conflated under 'sign' in our model). Johansson (2006) offers an explicit model, one concerned mainly with the evolution of grammar from stage (4) onwards. His model misses both stages (2) and (3). Finally, Dessalles (2006, 2008) comes closest to Table 1 using different terminology and without an explicit model. He has 'words' where we have 'signs' and '(non)commutativity' is never mentioned. Concepts like 'semantic embedding', CARC and CCLI are unique to our model, although there are sim-

[7] In honeybees' waggle dance, the direction of the dance encodes the direction to the food source relative to the sun's azimuth in the field, and the velocity of the dance encodes the distance to the food source (Michelsen, 1999). However, this does not imply semantic compositionality, as it is not clear whether honeybees interpret the direction and the velocity of the dance as the direction and the distance to the food source, respectively. It is possible that they interpret the whole dance as the location of the food source, i.e. they might not recognize the two elements in the dance as meaningful on their own.

ilarities between CARC and Dessalles' 'semantic synthesis' ability. Also, Dessalles (2006, p. 149; 2008, pp. 55, 61) presents (2) as a possibility (with references to Nowak et al.) rather than a necessary stage.

4. The evolution of syntactic compositionality: adaptiveness of the stages

Roughly, the correlates of the evolution of syntactic compositionality of language are the following: 1. The number of rules describing the set of signs increases. 2. The number of cues for distinct interpretations increases. 3. The ambiguity of interpretation decreases. Grammar implies full syntax, while stages (1)–(3) are necessary compositional prerequisites for grammar. In addition, stage (4) (grammar) is a positive criterion for defining language and a negative criterion for defining protolanguage. We point out that irrespective of whether the evolution of language was gradual (Hurford, 2012; Newmeyer, 1991; Pinker & Bloom, 1990) or catastrophic (Bickerton, 1995, 1998; Chomsky, 2010; Rosselló, Alba, Martin, & Borrega, 2012) there is no reason to single out **one** stage as protolanguage. Thus, stages (1)–(3) roughly correspond to what Bickerton and Jackendoff call protolanguage (Bickerton, 1990; Jackendoff, 1999, 2002). We assume that all stages in Table 1 are adaptive per se (otherwise it would not be clear why they should have evolved). The traits that contribute to fitness are far more likely to be selected for than those that do not. Still, it may be hard to see how could one benefit from free concatenation **before** the emergence of grammar (see Table 1).

Before we can answer this question, we have to make some assumptions about stages (1)–(3). It is logical to presume that in the beginning there were no distinct word types, and it is plausible that the first words met the condition that agents must have parallel non-verbal ways (e.g. pointing) to achieve goals of interactions (Steels et al., 2002). As the noun/verb distinction stipulates a primitive grammar and syntax, there was by definition no noun/verb distinction before grammar (i.e. in stages (1)–(3)). Further, as noun/verb is the most basic distinction among word types both comparatively and pragmatically, and the one that shows remarkable complementarity (Nowak & Krakauer, 1999; Sole, 2005), there were plausibly no distinct word types before the noun/verb distinction (Heine & Kuteva, 2002, 2007; Luuk, 2009). Stage (3) could contribute to fitness only insofar as it relied on constraints on interpretation, otherwise coherent interpretation could not have emerged. The constraints were provided by a relevance criterion. Depending on the context, different sets of relevance criteria might have been evoked, e.g., logical possibility, pragmatic or ontological feasibility, direct and/or inferential unexpectedness and/or emotionality (Dessalles, 2008), etc. However, having constraints on interpretation is not enough – minimally, coherent communication requires consistent and shared constraints on interpretation. Cultural constraints on linguistic interpretation (CCLI) generally satisfy these conditions. By CCLI we denote the pragmatic, logical and ontological constraints that are not imposed on a linguistic expression grammatically or lexically but are necessary to narrow down its interpretation. CCLI are enhanced by cooperation and small group size. Members of small groups and coalitions know each other well and face similar situations, but even then, the unambiguity of CCLI is limited. In order to maximize consistency and sharability, constraints on linguistic interpretation had to be externalized. This is precisely the motivation for syntax to evolve, and constitutes the transition from stage (3) to (4) (from commutative to noncommutative concatenation).

As for the putative concatenations at stage (3), relying solely on CCLI to be understood, consider the following example. *Wolf stone* or *stone wolf*, uttered by X to Y in the presence of wolves and stones, might be readily understood as a suggestion to throw stones at wolves, assuming that X and/or Y have behaved similarly before and have (roughly) the same interpretations for *stone* and *wolf*. Crucially, *Wolf stone* / *stone wolf* is interpreted as a compound expression by CCLI alone, and does not presuppose any syntactic constraints (i.e. grammar). But is speaking adaptive in a situation like this? It would have been more efficient if X just threw stones at wolves and Y imitated X. If the common goal is present in the actual environment, the collaborators need not focus on a joint representation of it before acting (Gardenfors, 2004). However, suppose that X has access to stones and Y does not. Then, if X did not start stoning wolves by himself, it would have made sense for Y to say *wolf stone*. Gärdenfors is arguably correct: the pragmatic aspects of language are the most fundamental from an evolutionary point of view. It is obvious that this kind of communication, though limited, could still contribute to fitness (cf. Jackendoff, 1999; Jackendoff & Pinker, 2005). Of course, verbal communication would be as likely in situations where immediate action is not required and participants have enough time to commune. Then, *stone wolf* or *wolf field* might inform X that Y stoned wolves or saw them on the field earlier. Similar examples can be found in, e.g., Bowie (2008) and Dessalles (2008).

Our conclusions about the utility of CCLI can be divided into two parts. 1. CCLI are sufficient for the interpretation of complex expressions only insofar as they are consistent and shared, two conditions that are enhanced by cooperation and small group size. 2. The interpretation of complex expressions at stage (3) relies on CCLI alone, as opposed to CCLI and grammar at stage (4).

5. Language, number, and numeral

For a long time, at least since the posing of Wallace's paradox,[8] it has been speculated that the mathematical capacity is an offshoot of the language faculty. According

[8] About 125 year ago, Alfred Russel Wallace posed the paradox that "the gigantic development of the mathematical capacity is wholly unexplained by the theory of natural selection, and must be due to some altogether distinct cause", if only because it remained unused (Chomsky, 2010).

to Chomsky (2010) and Hurford (1990) [1987]), number is derivative of language. It seems to be an established fact that exact arithmetic – and, hence, the cognition of **N** – is mainly dependent on language-specific representations (i.e. the verbal number concept – Dehaene, Spelke, Pinel, Stanescu, & Tsivkin, 1999; Nieder, 2005; Wiese, 2003). For example, exact calculation tasks are dependent on left inferior frontal activation that is also involved in verbal association tasks (Dehaene et al., 1999; Petersen, Fox, Posner, Mintun, & Raichle, 1988; Vandenberghe, Price, Wise, Josephs, & Frackowiak, 1996; Wagner et al., 1998). Similarly, aphasia following left-hemispheric brain damage tends to be associated with a selective impairment in exact arithmetic in contrast to approximate numerical judgments which depend on viso-spatial networks (for details, see Dehaene et al., 1999; cf. Monti, Parsons, & Osherson, 2012).

As both systems represent over serial channel, there are also certain structural similarities between language and the numeral system. Most importantly, the set {elements, concatenation, embedding} describes both systems. However, our numeral systems are structurally more complex than natural language, as they stipulate concatenation and embedding for each digit. In (unary, binary, decimal, hexadecimal or other – depending on the notation) point representation, numerical embedding can be depicted graphically by $[\ldots[x_3[x_2[x_1]]]].[y_1[y_2[y_3[\ldots]]]]$, where x's are integral and y's are fractional digits. In both systems, the elements are signs (i.e. form-meaning pairs) – meaningful linguistic units in language and numerals in the numeral system. Both numerical and semantic embedding are noncommutative: $[1[2]] \neq [2[1]]$ and $[run + s] \neq *s + run$. However, the constraints that stipulate numerical and semantic embedding are very different. In positional notation, a succession of digits reflects their magnitude, but there is no universal principle of succession of meaningful linguistic units. The universal magnitude constraint on concatenation stipulates numerical embedding, much like grammatical constraints on concatenation stipulate semantic embedding. Thus, in both systems, embedding is stipulated by constraints on concatenation. In sum, there is evidence of the same elementary cognitive operations underlying language, number, and the numeral system.

6. Discussion

Embedding and concatenation are the general rules of structuring – viz., those of inward and outward expansion, respectively. In models of language evolution, there has been only one proposal of the inward expansion antedating the outward one. This proposal, now largely dismissed (Bickerton, 2003; Johansson, 2008; Sundquist, 2012; Tallerman, 2007), is that of a holistic protolanguage by Wray (1998, 2000). Wray's proposal was that holistic utterances of protolanguage were, in the advent of syntax, fractured into distinct words. The main counterargument to this, supported by Johansson's (2008) calculation, is that the structure of the holistic utterances would have been too ambiguous to yield distinct form-meaning pairs (i.e. words) for the fractioning. Thus, the alternative hypothesis, that of the initial outward expansion by concatenation, would have to be true.

Both modern language and the numeral system have constraints on concatenation that stipulate noncommutative embedding (semantic and numerical embedding, respectively). However, the constraints themselves are different. The observed numeral systems obey the universal magnitude constraint, but there is no universal constraint on concatenation in language. Instead, linguistic concatenation is constrained by grammar, i.e. language-specific noncommutative concatenation. Thus, there is evidence of the same compositional capacity underlying language and the numeral system.

The structure, derivation and evolution of language is given by the sequence (elements, concatenation, embedding). This sequence is both derivational and evolutionary, as each member of the sequence has the one(s) to its left as its logical and evolutionary prerequisite(s). Arguably, the sequence is the general principle by which language is structured and evolved. Starting with a limited set of signs, it then expands the set, first by concatenating and, in later stages, also by embedding the signs. With the support of Jackendoff (1999), Nowak et al. (2000), Diessel and Tomasello (2005), Johansson (2006) and Dessalles (2006), we arrive at the following four-stage evolutionary scale of syntactic compositionality: (1) signs, (2) increased number of signs, (3) commutative concatenation of signs, (4) grammar (noncommutative concatenation of signs), resulting in semantic embedding (initially, words in phrases and sentences). The scale is hierarchical, i.e. at each stage the conditions stipulated by the previous stage(s) apply as well. We show how all these stages can be adaptive per se (which could explain why they evolved), and argue that CARC and CCLI are preconditions for maintaining stages (2) and (3), respectively. A principal trait of the scale is its scope: up to the emergence of grammar. Differently from e.g. Dessalles (2006), Jackendoff (1999), Johansson (2006), we do not model stages beyond (4). Implications for ontogeny should not be taken as granted but our model predicts that children's inventory of elementary verbal signs (not necessarily words, as children may confuse phrases with words) must grow to reach a certain size before the concatenation starts. The model also predicts a (possibly unstable) stage of commutative concatenation preceding the noncommutative one.

Acknowledgements

We thank James Hurford, Noam Chomsky, Michael Corballis, Haldur Õim, Kate Arnold, Kim Sterelny, Keith Stenning and the anonymous reviewers for their many helpful comments and suggestions. All remaining mistakes are our own. Erkki Luuk was supported by the target-financed theme No. 0180078s08, the National Programme for Estonian Language Technology project "Semantic analysis of simple sentences 2", the European Regional Development Fund through the Estonian Center of Excellence in Computer Science, EXCS, and the Alexander von Humboldt Foundation.

References

Arnold, K., & Zuberbuhler, K. (2006). Language evolution: Semantic combinations in primate calls. *Nature, 441*(7091), 303.

Ascenzi, A., Benvenuti, A., & Segre, A. G. (1997). On the paleopathologic findings exhibited by the late *Homo erectus* of Ceprano, Italy. *Human Evolution, 12*(3), 189–196.

Asfaw, B., Gilbert, W. H., Beyene, Y., Hart, W. K., Renne, P. R., WoldeGabriel, G., et al. (2002). Remains of *Homo erectus* from Bouri, Middle Awash, Ethiopia. *Nature, 416*, 317–320.

Bar-Yosef, O., & Belfer-Cohenb, A. (2001). From Africa to Eurasia — early dispersals. *Quaternary International, 75*(1), 19–28.

Bickerton, D. (1990). *Language and species*. Chicago: University of Chicago Press.

Bickerton, D. (1995). *Language and human behavior*. Seattle: University of Washington Press.

Bickerton, D. (2003). Symbol and structure: A comprehensive framework for language evolution. In M. H. Christiansen & S. Kirby (Eds.), *Language evolution: The states of the art* (pp. 77–93). Oxford: Oxford University Press.

Bickerton, D. (1998). Catastrophic evolution: The case for a single step from protolanguage to full human language. In J. R. Hurford, M. Studdert-Kennedy, & C. Knight (Eds.), *Approaches to the evolution of language: Social and cognitive bases* (pp. 341–358). Cambridge: Cambridge University Press.

Boë, L.-J., Heim, J.-L., Honda, K., Maeda, S., Badin, P., & Abry, C. (2007). The vocal tract of newborn humans and Neanderthals: Acoustic capabilities and consequences for the debate on the origin of language. A reply to Lieberman (2007a). *Journal of Phonetics, 35*, 564–581.

Bowie, J. (2008). Proto-discourse and the emergence of compositionality. *Interaction Studies, 9*(1), 18–33.

Budd, G. E. (2006). On the origin and evolution of major morphological characters. *Biological Reviews of the Cambridge Philosophical Society, 81*(4), 609–628.

Call, J., & Tomasello, M. (2008). Does the chimpanzee have a theory of mind? 30 years later. *Trends in Cognitive Sciences, 12*(5), 187–192.

Chomsky, N. (2010). Some simple evo-devo theses: How true might they be for language? In R. K. Larson, H. Yamakido, & V. Deprez (Eds.), *Evolution of human language: Biolinguistic perspectives* (pp. 45–61). Cambridge: Cambridge University Press.

Christiansen, M. H., & Chater, N. (2008). Language as shaped by the brain. *Behavioral and Brain Sciences, 31*, 489–558.

Christiansen, M. H., & Chater, N. (2003). Constituency and recursion in language. In M. A. Arbib (Ed.), *The handbook of brain theory and neural networks* (2nd ed., pp. 267–271). Cambridge, MA: MIT Press.

Christiansen, M. H., & Kirby, S. (2003). Language evolution: Consensus and controversies. *Trends in Cognitive Sciences, 7*(7), 300–307.

Correia, S. P., Dickinson, A., & Clayton, N. S. (2007). Western scrub-jays anticipate future needs independently of their current motivational state. *Current Biology, 17*(10), 856–861.

Dehaene, S., Spelke, E., Pinel, P., Stanescu, R., & Tsivkin, S. (1999). Sources of mathematical thinking: Behavioral and brain-imaging evidence. *Science, 284*(5416), 970–974.

Dessalles, J.-L. (2006). Du protolangage au langage: Modèle d'une transition. *Marges Linguistiques, 11*, 142–152.

Dessalles, J.-L. (2008). From metonymy to syntax in the communication of events. *Interaction Studies, 9*(1), 51–65.

Diessel, H., & Tomasello, M. (2005). A new look at the acquisition of relative clauses. *Language, 81*, 1–25.

Doumas, L. A. A., & Hummel, J. E. (2005). Approaches to modeling human mental representations: What works, what doesn't and why. In K. J. Holyoak & R. G. Morrison (Eds.), *The Cambridge handbook of thinking and reasoning* (pp. 73–91). New York: Cambridge University Press.

Fauconnier, G., & Turner, M. (2008). The origin of language as a product of the evolution of double-scope blending. *Behavioral and Brain Sciences, 31*(5), 520–521.

Fitch, W. T. (2000). The evolution of speech: A comparative review. *Trends in Cognitive Sciences, 4*(7), 258–267.

Gardenfors, P. (2004). Cooperation and the evolution of symbolic communication. In K. Oller & U. Griebe (Eds.), *Evolution of communication systems: A comparative approach* (pp. 237–256). Cambridge, MA: MIT Press.

Gardner, T. J., Naef, F., & Nottebohm, F. (2005). Freedom and rules: The acquisition and reprogramming of a bird's learned song. *Science, 308*(5724), 1046–1049.

Gibbons, A. (1998). In China, a handier *Homo erectus*. *Science, 279*(5357), 1636.

Givón, T. (1998). On the co-evolution of language, mind, and brain. *Evolution of Communication, 2*, 45–116.

Graffi, G. (2001). The emphasis on syntax in the early phase of European structuralism: Ries, Jespersen, Mathesius, Guillaume, Tesnière. In S. Auroux, E. F. K. Koerner, H.-J. Niederehe, & K. Versteegh (Eds.). *History of the language sciences/Geschichte der Sprachwissenschaften/Histoire des sciences du langage* (Vol. 2, pp. 1838–1847). Berlin – New York: De Gruyter.

Hauser, M. D., Chomsky, N., & Fitch, W. T. (2002). The faculty of language: What is it, who has it, and how did it evolve? *Science, 298*(5598), 1569–1579.

Heine, B., & Kuteva, T. (2007). *The genesis of grammar: A reconstruction*. New York: Oxford University Press.

Heine, B., & Kuteva, T. (2002). On the evolution of grammatical forms. In A. Wray (Ed.), *The transition to language* (pp. 376–397). Oxford: Oxford University Press.

Herman, L. M., Richards, D. G., & Wolz, J. P. (1984). Comprehension of sentences by bottlenosed dolphins. *Cognition, 16*(2), 129–219.

Hurford, J. R. (1990 [1987]). *Language and number: The emergence of a cognitive system*. Oxford: Basil Blackwell.

Hurford, J. R. (2004). Human uniqueness, learned symbols and recursive thought. *European Review, 12*(4), 551–565.

Hurford, J. R. (2012). Syntax evolved gradually. In T. C. Scott-Phillips, M. Tamariz, E. Cartmill, & J. R. Hurford (Eds.), *The evolution of language: Proceedings of the 9th international conference on the evolution of language* (pp. 165–172). Singapore: World Scientific.

Jackendoff, R. (1999). Possible stages in the evolution of the language capacity. *Trends in Cognitive Sciences, 3*(7), 272–279.

Jackendoff, R. (2002). *Foundations of language: Brain, meaning, grammar, evolution*. Oxford, New York: Oxford University Press.

Jackendoff, R., & Pinker, S. (2005). The nature of the language faculty and its implications for evolution of language (Reply to Fitch, Hauser, and Chomsky). *Cognition, 97*(2), 211–225.

Johansson, S. (2006). Working backwards from modern language to proto-grammar. In A. Cangelosi, A. D. M. Smith, & K. Smith (Eds.), *The evolution of language: Proceedings of the 6th international conference on the evolution of language (EVOLANG6)* (pp. 160–167). Singapore: World Scientific.

Johansson, S. (2008). Seeking compositionality in holistic protolanguage without substructure: Do counter-examples overwhelm the fractionation process? In A. D. M. Smith, K. Smith, & R. Ferrer-i-Cancho (Eds.), *The evolution of language: Proceedings of the 7th international conference (EVOLANG7)* (pp. 171–178). Singapore: World Scientific.

Kimbel, W. H., Johanson, D. C., & Rak, Y. (1997). Systematic assessment of a maxilla of *Homo* from Hadar, Ethiopia. *American Journal of Physical Anthropology, 103*, 235–262.

Klein, R. G. (2000). Archeology and the evolution of human behavior. *Evolutionary Anthropology, 9*(1), 17–36.

Knight, C. (2002). Language and revolutionary consciousness. In A. Wray (Ed.), *The transition to language* (pp. 138–160). Oxford: Oxford University Press.

Larick, R., Ciochon, R. L., Zaim, Y., Sudijono, Suminto, Rizal, Y., Aziz, F., et al. (2001). Early Pleistocene 40Ar/39Ar ages for Bapang Formation hominins, Central Jawa, Indonesia. *Proceedings of the National Academy of Sciences of the United States of America, 98*(9), 4866–4871.

Luuk, E. (2009). The noun/verb and predicate/argument structures. *Lingua, 119*(11), 1707–1727.

Luuk, E., & Luuk, H. (2008). Evolutionary framework for the language faculty. In A. D. M. Smith, K. Smith, & R. Ferrer-i-Cancho (Eds.), *The evolution of language: Proceedings of the 7th international conference (EVOLANG7)*. Singapore: World Scientific.

MacLarnon, A. M., & Hewitt, G. P. (1999). The evolution of human speech: The role of enhanced breathing control. *American Journal of Physical Anthropology, 109*, 341–363.

Martínez, I., Rosa, M., Arsuaga, J.-L., Jarabo, P., Quam, R., Lorenzo, C., et al. (2004). Auditory capacities in Middle Pleistocene humans from the Sierra de Atapuerca in Spain. *Proceedings of the National Academy of Sciences of the United States of America, 101*(27), 9976–9981.

Meyer, M. R. (2003). Vertebrae and language ability in early hominids. *PaleoAnthropology, 1*, 20–21.

Meyer, M. R., Lordkipanidze, D., & Vekua, A. (2006). Evidence for the anatomical capacity for spoken language in *Homo erectus*. *American Journal of Physical Anthropology Supplement, 42*, 130.

Michelsen, A. (1999). The dance language of honeybees: Recent findings and problems. In M. D. Hauser & M. Konishi (Eds.), *The design of animal communication* (pp. 111–131). Cambridge, MA: MIT Press.

Monti, M. M., Parsons, L. M., & Osherson, D. N. (2012). Thought beyond language: A neural dissociation of algebra and natural language. *Psychological Science, 23*(8), 914–922.

Morwood, M. J., O'Sullivan, P. B., Aziz, F., & Raza, A. (1998). Fission-track ages of stone tools and fossils on the east Indonesian island of Flores. *Nature, 392*, 173–176.

Newmeyer, F. J. (1991). Functional explanation in linguistics and the origins of language. *Language and Communication, 11*(1–2), 3–28.

Nieder, A. (2005). Counting on neurons: The neurobiology of numerical competence. *Nature Reviews Neuroscience, 6*(3), 177–190.

Nowak, M. A. (2000). Evolutionary biology of language. *Philosophical Transactions of the Royal Society of London B Biological Sciences, 355*(1403), 1615–1622.

Nowak, M. A., & Komarova, N. L. (2001). Towards an evolutionary theory of language. *Trends in Cognitive Sciences, 5*(7), 288–295.

Nowak, M. A., Komarova, N. L., & Niyogi, P. (2001). Evolution of universal grammar. *Science, 291*(5501), 114–118.

Nowak, M. A., & Krakauer, D. C. (1999). The evolution of language. *Proceedings of the National Academy of Sciences of the United States of America, 96*(14), 8028–8033.

Nowak, M. A., Plotkin, J. B., & Jansen, V. A. (2000). The evolution of syntactic communication. *Nature, 404*(6777), 495–498.

Osvath, M. (2009). Spontaneous planning for future stone throwing by a male chimpanzee. *Current Biology, 19*(5), R190–R191.

Penn, D. C., Holyoak, K. J., & Povinelli, D. J. (2008). Darwin's mistake: Explaining the discontinuity between human and nonhuman minds. *Behavioral and Brain Sciences, 31*(2), 109–130, discussion 130–178.

Penn, D. C., & Povinelli, D. J. (2007). On the lack of evidence that non-human animals possess anything remotely resembling a 'theory of mind'. *Philosophical Transactions of the Royal Society B: Biological Sciences, 362*(1480), 731–744.

Perner, J. (1991). *Understanding the representational mind*. Cambridge, MA: MIT Press.

Petersen, S. E., Fox, P. T., Posner, M. I., Mintun, M., & Raichle, M. E. (1988). Positron emission tomographic studies of the cortical anatomy of single-word processing. *Nature, 331*(6157), 585–589.

Pinker, S., & Bloom, P. (1990). Natural language and natural selection. *Behavioral and Brain Sciences, 13*(4), 707–784.

Premack, D., & Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *Behavioral and Brain Sciences, 1*(4), 515–526.

Rosselló, J., Alba, C., Martin, T., & Borrega, O. (2012). No words without syntax no syntax without words. In T. C. Scott-Phillips, M. Tamariz, E. Cartmill, & J. R. Hurford (Eds.), *The evolution of language: Proceedings of the 9th international conference on the evolution of language* (pp. 527–528). Singapore: World Scientific.

Savage-Rumbaugh, S., Shanker, S., & Taylor, T. (1998). *Apes, language, and the human mind*. New York: Oxford University Press.

Seyfarth, R., Cheney, D., & Bergman, T. J. (2005). Primate social cognition and the origins of language. *Trends in Cognitive Sciences, 9*(6), 264–266.

Sole, R. V. (2005). Syntax for free? *Nature, 434*, 289.

Spoor, F., Leakey, M. G., Gathogo, P. N., Brown, F. H., Antón, S. C., McDougal, I., et al. (2007). Implications of new early *Homo* fossils from Ileret, east of Lake Turkana, Kenya. *Nature, 448*, 688–691.

Steels, L., Kaplan, F., McIntyre, A., & Looveren, J. V. (2002). Crucial factors in the origins of word-meaning. In A. Wray (Ed.), *The transition to language* (pp. 252–271). Oxford: Oxford University Press.

Suddendorf, T., & Corballis, M. C. (2007). The evolution of foresight: What is mental time travel, and is it unique to humans? *Behavioral and Brain Sciences, 30*(3), 299–313.

Sundquist, J. D. (2012). Grammaticalization, light verbs, and the origins of complex language. In T. C. Scott-Phillips, M. Tamariz, E. Cartmill, & J. R. Hurford (Eds.), *The evolution of language: Proceedings of the 9th international conference on the evolution of language* (pp. 549–550). Singapore: World Scientific.

Szamado, S., & Szathmary, E. (2006). Selective scenarios for the emergence of natural language. *Trends in Ecology & Evolution, 21*(10), 555–561.

Tallerman, M. (2007). Did our ancestors speak a holistic protolanguage? *Lingua, 117*, 579–604.

Tobias, P. V. (1998). Evidence for the early beginnings of spoken language. *Cambridge Archaeological Journal, 8*(1), 72–78.

Tomasello, M., Call, J., & Hare, B. (2003). Chimpanzees understand psychological states – The question is which ones and to what extent. *Trends in Cognitive Sciences, 7*(4), 153–156.

Tulving, E. (2005). Episodic memory and autonoesis: Uniquely human? In H. S. Terrace & J. Metcalfe (Eds.), *The missing link in cognition: Origins of self-reflective consciousness* (pp. 3–56). Oxford: Oxford University Press.

Ujhelyi, M. (1998). Long call structure in apes as a possible precursor for language. In J. R. Hurford, M. Studdert-Kennedy, & C. Knight (Eds.), *Approaches to the evolution of language: Social and cognitive bases*. Cambridge, UK: Cambridge University Press.

Vandenberghe, R., Price, C., Wise, R., Josephs, O., & Frackowiak, R. S. (1996). Functional anatomy of a common semantic system for words and pictures. *Nature, 383*(6597), 254–256.

von Frisch, K., & Lindauer, M. (1996). The "language" and orientation of the honey bee. In L. D. Houck & L. C. Drickamer (Eds.), *Foundations of animal behavior: Classic papers with commentaries* (pp. 539–552). Chicago: University of Chicago Press.

Wagner, A. D., Schacter, D. L., Rotte, M., Koutstaal, W., Maril, A., Dale, A. M., et al. (1998). Building memories: Remembering and forgetting of verbal experiences as predicted by brain activity. *Science, 281*(5380), 1188–1191.

Waters, M. R., Forman, S. L., & Pierson, J. M. (1997). Diring Yuriakh: A lower paleolithic site in Central Siberia. *Science, 275*, 1281–1284.

Wiese, H. (2003). Iconic and non-iconic stages in number development: The role of language. *Trends in Cognitive Sciences, 7*(9), 385–390.

Wray, A. (1998). Protolanguage as a holistic system for social interaction. *Language and Communication, 18*(1), 47–67.

Wray, A. (2000). Holistic utterances in protolanguage: The link from primates to humans. In C. Knight, J. R. Hurford, & M. Studdert-Kennedy (Eds.), *The evolutionary emergence of language: Social* *function and the origins of linguistic form* (pp. 285–302). Cambridge, UK: Cambridge University Press.

Wynn, T. (1991). Tools, grammar and the archaeology of cognition. *Cambridge Archaeological Journal, 1*(2), 191–206.

Wynn, T. (1998). Did *Homo erectus* speak? *Cambridge Archaeological Journal, 8*(1), 78–81.

Wynn, T. (2002). Archaeology and cognitive evolution. *Behavioral and Brain Sciences, 25*, 389–438.